1 Details for the Composite Likelihood Implementation

This vignette covers computational details about the implementation. It is provided for interested users. The vignette largely follows Cederkvist et al. (2018) but we point out a few additional ways to speed up the computation and how to make an implementation.

We have n clusters with each having m_i members. Let T^*_{ij}, ϵ^*_{ij}, and C_{ij} be the failure time, the cause of failure, and the censoring time of individual j in cluster i, respectively. There are K competing risks and, thus, $\epsilon^*_{ij} \in \{1, \dots, K\}$. Let $T_{ij} = \min(T^*_{ij}, C_{ij})$ be the observed time, $\Lambda_{ij} = I(T^*_{ij} \leq C_{ij})$ be an event indicator and $\epsilon_{ij} = \Lambda_{ij}\epsilon^*_{ij} \in \{0, 1, \dots, K\}$ be the observed failure cause.

Each cluster has a cluster specific random effect given by

$$\begin{pmatrix} \boldsymbol{U}_i \\ \boldsymbol{\eta}_i \end{pmatrix} \sim N^{(2K)}(\mathbf{0}, \mathbf{\Sigma})$$

where $\boldsymbol{U}_i, \boldsymbol{\eta}_i \in \mathbb{R}^K$ and $\sim N^{(V)}(\mathbf{0}, \mathbf{\Psi})$ indicates that a random follows a V dimensional multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Psi}$. The conditional density of observing event k at the observed time t for individual j in cluster i is

$$F_{kij}(t \mid \boldsymbol{u}_i, \boldsymbol{\eta}_i) = \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\Phi(-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k - \eta_{ik})$$

where

$$\pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i) = \frac{\exp(\boldsymbol{z}_{ij}^\top \boldsymbol{\beta}_k + u_{ik})}{1 + \sum_{l=1}^K \exp(\boldsymbol{z}_{ij}^\top \boldsymbol{\beta}_l + u_{il})}$$

and $\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k$ is monotonically decreasing. Thus, the conditional survival probability is

$$1 - \sum_{k=1}^K F_{kij}(t \mid \boldsymbol{u}_i, \boldsymbol{\eta}_i).$$

We allow for $-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k = \infty$ such that $F_{kij}(t \mid \boldsymbol{u}_i, \boldsymbol{\eta}_i) = \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)$. In the case where this is true for all $k \in \{1, \dots, K\}$, the conditional survival probability simplifies to

$$\pi_0(\boldsymbol{z}_{ij}, \boldsymbol{u}_i) = \frac{1}{1 + \sum_{l=1}^K \exp(\boldsymbol{z}_{ij}^\top \boldsymbol{\beta}_l + u_{il})}.$$

This can be used to greatly simplify the computations.

Composite Likelihood

The model is estimated with pairwise composite likelihood. This leads to three types of the log composite likelihood terms. The first type is when both j' and j are observed with failure cause k' and k at time t' and t. The term is given by

$$\begin{aligned}\log(-\boldsymbol{x}'_{ij}(t)^\top \boldsymbol{\gamma}_k) + \log(-\boldsymbol{x}'_{ij'}(t')^\top \boldsymbol{\gamma}_{k'}) + \log \int\int & \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\phi(-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k - \eta_{ik}) \\ & \cdot \pi_{k'}(\boldsymbol{z}_{ij'}, \boldsymbol{u}_i)\phi(-\boldsymbol{x}_{ij'}(t')^\top \boldsymbol{\gamma}_{k'} - \eta_{ik'})\phi^{(2K)}\left(\begin{pmatrix} \boldsymbol{u}_i \\ \boldsymbol{\eta}_i \end{pmatrix}, \mathbf{0}, \mathbf{\Sigma}\right) \mathrm{d}\boldsymbol{u}_i \, \mathrm{d}\boldsymbol{\eta}_i\end{aligned}$$

where the derivatives are w.r.t. time and $\phi^{(2K)}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{\Sigma})$ is the $2K$ dimensional multivariate normal distribution's density with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$. We can use that

$$\begin{aligned}&\phi^{(2)}\left(\begin{pmatrix} -\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k \\ -\boldsymbol{x}_{ij'}(t')^\top \boldsymbol{\gamma}_{k'} \end{pmatrix}; \mathbf{V}\begin{pmatrix} \boldsymbol{u}_i \\ \boldsymbol{\eta}_i \end{pmatrix}, \mathbf{I}_2\right)\phi^{(2K)}\left(\begin{pmatrix} \boldsymbol{u}_i \\ \boldsymbol{\eta}_i \end{pmatrix}; \mathbf{0}, \mathbf{\Sigma}\right) \\ &\qquad = \phi^{(2)}\left(\begin{pmatrix} -\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k \\ -\boldsymbol{x}_{ij'}(t')^\top \boldsymbol{\gamma}_{k'} \end{pmatrix}; \mathbf{0}, \mathbf{I}_2 + \mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top\right)\phi^{(2K)}\left(\begin{pmatrix} \boldsymbol{u}_i \\ \boldsymbol{\eta}_i \end{pmatrix}; \mathbf{M}\mathbf{V}^\top \begin{pmatrix} -\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k \\ -\boldsymbol{x}_{ij'}(t')^\top \boldsymbol{\gamma}_{k'} \end{pmatrix}, \mathbf{M}\right)\end{aligned}$$

where $\mathbf{I}_l$ is the l dimensional identity matrix, $\mathbf{V}$ is a matrix containing zeros except for a one in the $K+k$th entry in the first row and the $K+k'$th entry in the second row and $\mathbf{M} = (\mathbf{V}^\top\mathbf{V} + \mathbf{\Sigma}^{-1})^{-1}$. Thus, we can re-write the log composite likelihood term as

$$\begin{aligned}\log(-\boldsymbol{x}_{ij}'(t)^\top\boldsymbol{\gamma}_k) &+ \log(-\boldsymbol{x}_{ij'}'(t')^\top\boldsymbol{\gamma}_{k'}) + \log\phi^{(2)}\left(\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\ -\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix};\mathbf{0},\mathbf{I}_2+\mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top\right)\\ &+\log\int\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\ \pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\ \boldsymbol{\eta}_i\end{pmatrix};\mathbf{M}\mathbf{V}^\top\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\ -\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix},\mathbf{M}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i\end{aligned}$$

which further simplifies to

$$\begin{aligned}\log(-\boldsymbol{x}_{ij}'(t)^\top\boldsymbol{\gamma}_k) &+ \log(-\boldsymbol{x}_{ij'}'(t')^\top\boldsymbol{\gamma}_{k'}) + \log\phi^{(2)}\left(\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\ -\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix};\mathbf{0},\mathbf{I}_2+\mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top\right)\\ &+\log\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\ \pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i;\mathbf{M}_{1:K,\cdot}\mathbf{V}^\top\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\ -\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix},\mathbf{M}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i\end{aligned}$$

where $\mathbf{M}_{1:l,1:l'}$ is the first $l\times l'$ block of $\mathbf{M}$ and a '$\cdot$' denotes all rows or columns. The problem can be standardized to working with fixed values $\boldsymbol{a}_{ijl} = -\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_l$'s and $\boldsymbol{b}_{ijl} = \boldsymbol{z}_{ij}^\top\boldsymbol{\beta}_l$'s and matrix $\mathbf{M}$ and computing the derivatives w.r.t. these quantities, $\boldsymbol{a}_{ijl}$'s, $\boldsymbol{b}_{ijl}$'s and $\mathbf{M}$. The chain rule can then be applied to get the derivatives w.r.t. the $\boldsymbol{\gamma}_l$'s, $\boldsymbol{\beta}_l$'s, and $\mathbf{\Sigma}$. This is computationally very fast. The sparsity of $\mathbf{V}$ can also be used to simplify the expression above.

With one censored individual j' and observed outcome for j with failure cause k, the log composite likelihood term is

$$\begin{aligned}\log(-\boldsymbol{x}_{ij}'(t)^\top\boldsymbol{\gamma}_k) + \log\int\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\phi(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k-\eta_{ik})\left(1-\sum_{k'=1}^K F_{k'ij'}(t'\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)&\\ \cdot\,\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\ \boldsymbol{\eta}_i\end{pmatrix},\mathbf{0},\mathbf{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i.&\end{aligned}$$

Again, we can turn around the conditioning to get

$$\begin{aligned}\log(-\boldsymbol{x}_{ij}'(t)^\top\boldsymbol{\gamma}_k) &+ \log\phi\left(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k,0,1+\boldsymbol{v}^\top\mathbf{\Sigma}\boldsymbol{v}\right)\\ &+\log\int\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\left(1-\sum_{k'=1}^K F_{k'ij'}(t'\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\ \boldsymbol{\eta}_i\end{pmatrix};\mathbf{M}\boldsymbol{v}(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k),\mathbf{M}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i\end{aligned}$$

where $\mathbf{M} = (\boldsymbol{v}\boldsymbol{v}^\top + \mathbf{\Sigma}^{-1})^{-1}$ and $\boldsymbol{v}$ is a $2K$ vector with zeros except at the $K+k$th entry which is one. The result is $K+1$ intractable integrals of dimension K. The first integral is

$$\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i;\mathbf{M}_{1:K,\cdot}\boldsymbol{v}(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k),\mathbf{M}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i.$$

The remaining K integrals are of the form

$$\begin{aligned}\int\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)&\pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\Phi(-\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}-\eta_{ik'})\\ &\cdot\,\phi^{(K+1)}\left(\begin{pmatrix}\boldsymbol{u}_i\\ \eta_{ik'}\end{pmatrix};\begin{pmatrix}\boldsymbol{\mu}_{1:K}\\ \mu_{K+k'}\end{pmatrix},\begin{pmatrix}\mathbf{M}_{1:K,1:K},\mathbf{M}_{1:K,K+k'}\\ \mathbf{M}_{K+k',1:K},\mathbf{M}_{K+k',K+k'}\end{pmatrix}\right)\mathrm{d}\eta_{ik'}\,\mathrm{d}\boldsymbol{u}_i\\ &=\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\Phi\left(\frac{-\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}-g(\boldsymbol{u}_i)}{s}\right)\phi^{(K)}\left(\boldsymbol{u}_i;\boldsymbol{\mu}_{1:K},\mathbf{M}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i\end{aligned}$$

where $\boldsymbol{\mu} = \mathbf{M}\boldsymbol{v}(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k)$, $g(\boldsymbol{u}_i) = \mu_{K+k'} + \mathbf{M}_{K+k',1:K}\mathbf{M}_{1:K,1:K}^{-1}(\boldsymbol{u}_i-\boldsymbol{\mu}_{1:K})$ and

$$s^2 = 1 + \mathbf{M}_{K+k',K+k'} - \mathbf{M}_{K+k',1:K}\mathbf{M}_{1:K,1:K}^{-1}\mathbf{M}_{1:K,K+k'}.$$

Note that if $-\boldsymbol{x}_{ij'}(t')^\top\gamma_k = \infty$ for all $k \in \{1,\dots,K\}$, then log composite likelihood term is

$$\log(-\boldsymbol{x}_{ij}'(t)^\top\boldsymbol{\gamma}_k) + \log\phi\left(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k, 0, 1 + \boldsymbol{v}^\top\boldsymbol{\Sigma}\boldsymbol{v}\right) \\ + \log\int \pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\pi_0(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i;\mathbf{M}_{1:K\cdot}\boldsymbol{v}(-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k),\mathbf{M}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i.$$

This is computationally easier to evaluate.

Finally, we get the following log composite likelihood term if both individuals are censored

$$\log\left(\int\int\left(1-\sum_{k=1}^K F_{kij}(t\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)\left(1-\sum_{k=1}^K F_{kij'}(t'\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\\boldsymbol{\eta}_i\end{pmatrix},\mathbf{0},\boldsymbol{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i\right) \\ = \log\left(\int\int\left(\cdots+\sum_{k=1}^K\sum_{k'=1}^K F_{kij}(t\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)F_{k'ij'}(t'\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\\boldsymbol{\eta}_i\end{pmatrix},\mathbf{0},\boldsymbol{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i\right)$$

where we cover how to compute the first $2K$ integrals that are not shown in Section 1.1. The final K^2 integrals are of the form:

$$\int\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\Phi^{(2)}\left(\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\-\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix}-\mathbf{V}\boldsymbol{\eta}_i\right)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\\boldsymbol{\eta}_i\end{pmatrix},\mathbf{0},\boldsymbol{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i$$

where $\Phi^{(2)}$ is the bivariate standard normal CDF integrated over the rectangle from minus infinity to the passed upper bounds and $\mathbf{V}$ is matrix with zeros except for a one in the kth entry in the first row and k'th entry in the second row. We can re-write this as

$$\int\pi_k(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\pi_{k'}(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\Phi^{(2)}\left(\begin{pmatrix}-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k\\-\boldsymbol{x}_{ij'}(t')^\top\boldsymbol{\gamma}_{k'}\end{pmatrix}-\mathbf{V}\boldsymbol{g}(\boldsymbol{u}_i);\mathbf{0},\mathbf{I}_2+\mathbf{V}\mathbf{M}\mathbf{V}^\top\right)\phi^{(K)}\left(\boldsymbol{u}_i;\mathbf{0},\boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i$$

where $\boldsymbol{g}(\boldsymbol{u}_i) = \boldsymbol{\Sigma}_{(K+1):2K,1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{u}_i$, $\mathbf{M} = \boldsymbol{\Sigma}_{(1+K):2K,(1+K):2K} - \boldsymbol{\Sigma}_{(1+K):2K,1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{\Sigma}_{1:K,(1+K):2K}$ and $\Phi^{(2)}$ is the CDF of a bivariate normal distribution with the specified mean and covariance matrix. The CDF can be solved efficiently using one dimensional quadrature using the method mentioned in Genz (2004). Nevertheless, this additional application of quadrature makes these integrals computationally more demanding than the other integrals we have shown till now.

If one individual j' has $-\boldsymbol{x}_{ij'}(t')^\top\gamma_k = \infty$ for all $k \in \{1,\dots,K\}$, then the log composite likelihood term is

$$\log\left(\int\int\left(1-\sum_{k=1}^K F_{kij}(t\mid\boldsymbol{u}_i,\boldsymbol{\eta}_i)\right)\pi_0(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i\\\boldsymbol{\eta}_i\end{pmatrix},\mathbf{0},\boldsymbol{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i\right)$$

This lead to integrals of the form

$$\int\pi_0(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i,\mathbf{0},\boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i$$

and

$$\int\pi_0(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\pi_k(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\Phi\left(\frac{-\boldsymbol{x}_{ij}(t)^\top\boldsymbol{\gamma}_k - g(\boldsymbol{u}_i)}{s}\right)\phi^{(K)}\left(\boldsymbol{u}_i,\mathbf{0},\boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i$$

where $g(\boldsymbol{u}_i) = \boldsymbol{\Sigma}_{K+k,1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{u}_i$ and

$$s^2 = 1 + \boldsymbol{\Sigma}_{K+k,K+k} - \boldsymbol{\Sigma}_{K+k,1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{\Sigma}_{1:K,K+k}.$$

If also $-\boldsymbol{x}_{ij}(t)^\top\gamma_k = \infty$ for all $k \in \{1,\dots,K\}$, then the log composite likelihood term is

$$\log\left(\int\pi_0(\boldsymbol{z}_{ij},\boldsymbol{u}_i)\pi_0(\boldsymbol{z}_{ij'},\boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i,\mathbf{0},\boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i\right).$$

To summarize, we have to, at-worst, compute

1. One K dimensional integral when both individuals are observed.

2. K integrals of dimension K when one individual is censored.

3. $2K$ integrals of dimension K and K^2 integrals of dimension $K + 1$ when both individuals are censored.

Preliminary experiments using `https://github.com/boennecd/ghq-cpp/tree/main/ghqCpp` shows that we can compute each of the K dimensional integrals in about ten microseconds or less when $K = 2$ with adaptive Gauss-Hermite quadrature. It takes a bit longer when one also has to use the method mentioned in Genz (2004). The cost of all other computations are negligible.

1.1 Singleton Observations

We may have clusters with only one observation. In this case, the log composite likelihood if the individual is observed is

$$\log(-\boldsymbol{x}_{ij}'(t)^\top \boldsymbol{\gamma}_k) + \log \int\int \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\phi(-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k - \eta_{ik})\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i \\ \boldsymbol{\eta}_i\end{pmatrix}, \boldsymbol{0}, \boldsymbol{\Sigma}\right) \mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i.$$

Turning the conditioning around, we have

$$\log(-\boldsymbol{x}_{ij}'(t)^\top \boldsymbol{\gamma}_k) + \log\phi\left(-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k, 0, 1 + \boldsymbol{v}^\top \boldsymbol{\Sigma}\boldsymbol{v}\right) + \log\int \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i; \boldsymbol{\mu}_{1:K}, \mathbf{M}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i$$

where $\boldsymbol{v}$ is a vector with zero expect in the $K + k$'th entry, $\mathbf{M} = (\boldsymbol{v}\boldsymbol{v}^\top + \boldsymbol{\Sigma}^{-1})^{-1}$ and $\boldsymbol{\mu} = \mathbf{M}\boldsymbol{v}(-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k)$.

The log composite likelihood term if the individual is censored is

$$\log\int\int\left(1 - \sum_{k=1}^{K} F_{kij}(t \mid \boldsymbol{u}_i, \boldsymbol{\eta}_i)\right)\phi^{(2K)}\left(\begin{pmatrix}\boldsymbol{u}_i \\ \boldsymbol{\eta}_i\end{pmatrix}, \boldsymbol{0}, \boldsymbol{\Sigma}\right)\mathrm{d}\boldsymbol{u}_i\,\mathrm{d}\boldsymbol{\eta}_i.$$

Leading to K terms of the form

$$\int \pi_k(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\Phi\left(\frac{-\boldsymbol{x}_{ij}(t)^\top \boldsymbol{\gamma}_k - g(\boldsymbol{u}_i)}{s}\right)\phi^{(K)}\left(\boldsymbol{u}_i; \boldsymbol{0}, \boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i$$

where $g(\boldsymbol{u}_i) = \boldsymbol{\Sigma}_{K+k',1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{u}_i$ and

$$s^2 = 1 + \boldsymbol{\Sigma}_{K+k',K+k'} - \boldsymbol{\Sigma}_{K+k',1:K}\boldsymbol{\Sigma}_{1:K,1:K}^{-1}\boldsymbol{\Sigma}_{1:K,K+k'}.$$

If $-\boldsymbol{x}_{ij}(t)^\top \gamma_k = \infty$ for all $k \in \{1, \dots, K\}$, then we can instead compute

$$\int \pi_0(\boldsymbol{z}_{ij}, \boldsymbol{u}_i)\phi^{(K)}\left(\boldsymbol{u}_i; \boldsymbol{0}, \boldsymbol{\Sigma}_{1:K,1:K}\right)\mathrm{d}\boldsymbol{u}_i.$$

Again, this is computationally easier to work with.

References

Cederkvist, L., Holst, K. K., Andersen, K. K., and Scheike, T. H. (2018). Modeling the cumulative incidence function of multivariate competing risks data allowing for within-cluster dependence of risk and timing. *Biostatistics*, 20(2):199–217.

Genz, A. (2004). Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing*, 14(3):251–260.